Is a more liberal approach to conservatism needed in forecasting?

Paul Goodwin, The School of Management, University of Bath, Bath, BA2 7AY,

United Kingdom

Email: p.goodwin@bath.ac.uk

Submission: June 2014
Revision: December 2014
Acceptance: January 2015

Is a more liberal approach to conservatism needed in forecasting?

ABSTRACT

I discuss evidence that supports several of the principles put forward in the paper by Armstrong, Green, and Graefe (AGG), but argue that the packaging of these principles as a single "golden rule"' and the use of the term "conservative" may lead to misunderstandings. Additional work should be carried out to investigate the extent to which these principles should be applied to probability and interval forecasts. Finally, good reasons may support why "rational" forecasters behave in ways that are inconsistent with the guidelines AGG provide in their golden-rule checklist.

Keywords: forecasting, conservatism, probability forecasts, prediction intervals.

1.1 Evidence for the guidelines

Armstrong, Green, and Graefe (referred to as AGG below) offer much good advice and many of the points they make chime with my observations of forecasting practices in supply-chain companies.

"Claims that things are different now should be met with skeptism" accords with the common experience of seeing forecasters summarily dumping their computer-based forecasts in favour of unsupported judgment because each new month is regarded as a special case. Indeed, the tendency to neglect dull base-rates (e.g. statistical forecasts) in favour of more colourful, but less reliable, anecdotal information or rumour or sentiment has been reported by psychologists for forty years (Tversky and Kahneman 1974). Also, Kahneman and Lovallo (1993) distinguish between the "inside view", where people focus on the specific details of the situation they are forecasting, and the "outside view" where they downplay these details and look at the "class of cases" that are similar to the current case. The tendency to adopt the inside view fosters a perception that the situation that is being forecast is unique and that past observations are irrelevant.

Similarly, AGG's advice to "conceal the purpose of the forecast", where this is possible, would prevent forecasts being turned into decisions. As such, the forecast is more likely to remain as an honest expectation of what will happen in the future, rather than a choice intended to achieve an objective, such as minimizing costs, maximizing customer goodwill, or maximising the kudos of a particular section of an organization for political reasons (Goodwin, 1998, pp. 95-6). Blurring the distinction between forecasts and choice can lead to confusion amongst forecast users and it can cause problems when a decision, presented as a forecast, is used in another part of the

organization where other objectives apply (Fildes, Goodwin, Lawrence, and Nikolopoulos 2009).

I have also come across situations in companies where forecasters judgmentally select "a particular starting point [in time] for estimating a time-series forecasting model …. [allowing them] to make forecasts that support their prior beliefs" (section 3.1). In a pharmaceutical company forecasters praised their software because of the ease with which they could make these selections. But, conveniently, they referred to the resulting forecasts as the "system" forecasts and attributed any errors to the deficiencies of the software (Brown, Goodwin, and Fildes, 2011). Forecasting the effect of special events independently of the statistical forecasts also makes sense, though I have yet to see this idea tested or applied in practice.

2.0 Conservatism and nuanced universality

This commentary has undergone several revisions as AGG took on board some of the earlier points that I made in subsequent revisions of their paper. However, the paper's "headlines" give the impression that the diverse and often complex findings of forecasting research can be represented as a [single] "golden rule" with an associated simple and potentially misleading label: conservatism. I would have far fewer problems if the authors had referred to what their paper really presents which are multiple evidence-based rules applicable to different forecasting contexts. After all, who can argue with that, as long as the authors' interpretation of the evidence is accurate? Our raison d'être as forecasting researchers is to seek out evidence to guide and improve forecasting practice. A paper that organises and codifies this evidence is to be welcomed.

Take the label "conservatism". The Merrian-Webster dictionary defines this, inter alia, as a "dislike of change or new ideas in a particular area". Dictionary.com defines "conservative" as "cautiously moderate or purposefully low". My own initial impression was that the authors were advocating a universal emphasis on naïve or "no change" forecasts. This turns out not to be the case and their recommendations are much richer than the term implies. For example, AGG state that: "The no-change model is not always conservative. There are many cases where cumulative knowledge calls for change" –so why use the term conservative? It can only distract from, and confuse, the paper's potentially valuable underlying messages.

In some cases the rules in the Golden rule checklist would benefit from being more nuanced or specific. For example, 1.1.1 in the checklist states that forecasters should "decompose to best use knowledge and information and judgment". However, a study by MacGregor and Armstrong (1994) reports that multiplicative decomposition "improved accuracy only when the situation involved uncertain and extreme quantities". In other situations accuracy was reduced. Similarly, after a review of the evidence, Macgregor (2001) suggests that we should "use decomposition when uncertainty is high; otherwise use global or holistic estimation". To be fair, the checklist includes guideline 1.1.2: "Select evidence-based methods validated for the situation", but this should surely be an overarching guideline as it interacts with, and can therefore invalidate, other guidelines, such as 1.1.1, in particular situations.

3.0 Conservatism and probability forecasts

In addition to point forecasts, some measure of uncertainty or risk is often useful. The Italian earthquake forecasters, mentioned in the AGG paper, were jailed because they did not attach probabilities to their forecasts and simply conveyed the most

probable outcome, thereby giving the impression that there was no risk (Mazzotti 2013). Similarly, the UK Meteorological Office was recently criticised when it produced forecasts that "slightly favoured a drier than average" April 2012 (Gray 2013). It turned out to be the wettest April on record. However, the office had not publicised its internal forecast, which allowed for a 10% to 15% probability that the weather would be very wet. Good practice would surely encourage the greater use of probability forecasts and prediction intervals, but the paper gives us little guidance here and it seems that this is an area that would be worthy of future research.

Is being conservative about uncertainty also important? In studies of how people update probability forecasts in the light of new information, conservatism is usually considered to be a bad thing (Phillips and Edwards 1966). Here the term is used to describe a reluctance to change one's probability estimates when reliable and highly diagnostic information is received –in short a tendency to discount good evidence. Bayes theorem provides the normative approach to the updating of probability forecasts and the result may be frequent changes of forecasts as each new item of information is received. This normative approach may be consistent with the AGG's notion of conservatism, but once again the term is unhelpful and clearly there is a danger of it being used to represent very different forecasting strategies.

When it comes to prediction intervals, or forecasts in the form of probability distributions, how should the golden rule be applied? Indeed, does it have any direct implications, apart from providing guidance on how to estimate the central tendency of the associated probability distribution? The usual problem with these types of forecasts, irrespective of whether they are produced through statistical methods or judgment, is that they have ranges that are too narrow so that they enclose outcomes less frequently than they should according to their stated probabilities. Are wider

ranges consistent with conservatism? My guess is that they would be even though such ranges would acknowledge the possibility of large changes from the current situation, and even greater changes than might have been observed in the historical data series. The rules which state "specify multiple hypotheses and methods" (1.2.2), "Ask judges to write reasons for and against the forecasts"(2.3), "combine independent forecasts from judges (2.6)" and "combine forecasts from diverse evidence-based methods (5)" would all appear to encourage wider ranges.

4.0 Departing from the checklist

Finally, the checklist provided is intended to lead to improved forecasts if it is adhered to, but decisions regarding forecasting methods usually involve multiple objectives. Trade-offs have to be made between the cost of forecast errors, the cost of producing forecasts, the time to produce forecasts, and the need to make the forecasts credible. Methods like decomposition can be time consuming, employing heterogeneous experts can be expensive, and the use of very simple methods may lack credibility even if they are generally accurate. On that basis there may be good reasons why "rational" forecasters behave in ways that are not consistent with the checklist. Often it is reasonable to sacrifice some accuracy in order to achieve other objectives such as acceptability. Imposing penalties on forecasters for non-adherence to the list, as suggested, would therefore be inappropriate.

Nevertheless, overall the paper is stimulating and it serves a useful purpose by challenging many prevailing views. While one may not agree with the basis of all the challenges, the debates that they encourage should be to the benefit of forecasting research and practice.

References

Brown, C., Goodwin, P., and Fildes, R. Between success and failure: a forecasting system in use. UK Academy for Information Systems Conference Proceedings 2011; Paper 9. http://aisel.aisnet.org/ukais2011/9

Fildes, R., Goodwin P., Lawrence M., Nikolopoulos K. Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. International Journal of Forecasting 2009; 25: 3-23.

Goodwin, P. Enhancing judgmental sales forecasting: the role of laboratory research. In: G Wright and P. Goodwin editors, Forecasting with Judgment. Chichester: Wiley, 1998. pp. 91-111.

Gray, L. Met Office apologises for warning of 'dry spell' before wettest April on record. The Telegraph 2013: 29 March.

Kahneman, D., Lovallo, D. Timid choices and bold forecasts. A cognitive perspective on risk taking. Management Science 1993; 39: 17-31.

Mazzotti, M. Seismic shift. Times Higher Education, 2013; 2121: 38-43.

MacGregor, D. Decomposition for judgmental forecasting and estimation. In: J. S. Armstrong editor, Principles of Forecasting: A Handbook for Researchers and Practitioners. New York: Springer, 2001. pp. 107–123.

Macgregor, D.G., Armstrong, J.S. Judgmental decomposition: when does it work? International Journal of Forecasting 1994; 10: 495-506.

Phillips, L.D., Edwards, W. Conservatism in simple probability inference tasks. Journal of Experimental Psychology 1966; 72: 346-357.

Tversky, A., Kahneman, D. Judgment under uncertainty: Heuristics and biases. Science 1974; 185: 1124-1131.